Variance V.S. Bias

Mathematically:

Let the variable we are trying to predict as Y and other covariates as X. We assume there is a relationship between the two such that

$$Y = f(X) + e$$

where e is the error term and it's normally distributed with a mean of 0.

We will make a model $\hat{f}(X)$ of $f(X)$ using linear regression or any other modeling technique.

So, the expected squared error at a point X is

$$Err(x) = E\left[(Y - \hat{f}(x))^2 \right]$$

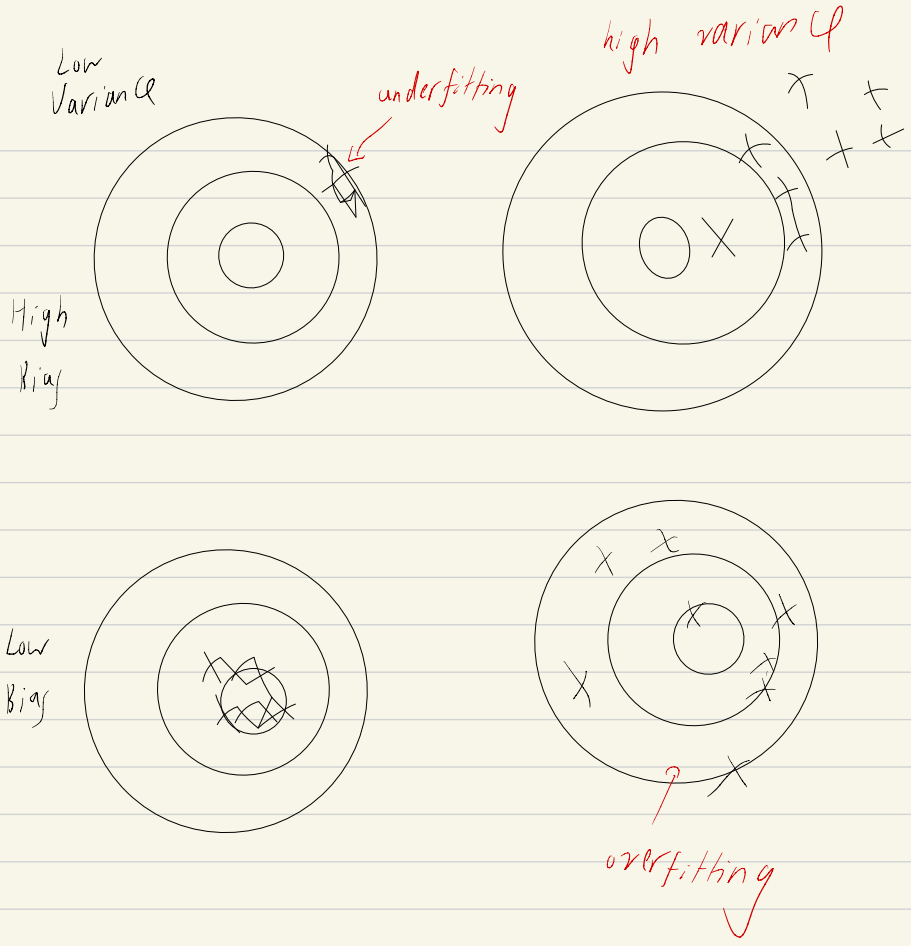
The $Err(x)$ can be further decomposed as

$$Err(x) = \left[E[\hat{f}(x)] - f(x) \right]^2 + E\left[\left(\hat{f}(x) - E[\hat{f}(x)] \right)^2 \right] + \sigma_e^2$$

$$Err(x) = Bias^2 + Variance + Irreducible\ Error.$$

Irreducible error is the error that can't be reduced by creating good models. It is a measure of the amount of noise in our data.

Here it is important to understand that no matter how good we make our model, our d



Why is Bias Variance Tradeoff?

If our model is too simple and has very few parameters then it may have high bias and low variance. On the other hand if our model has large number of parameters then it's going to have high variance and low bias. So we need to find the right/good balance without overfitting and underfitting the data.

This tradeoff in complexity is why there is a tradeoff between bias and variance. An algorithm